



Medical Drive, 117597, SINGAPORE  
Location/Qualifiers  
FEATURES  
source  
1..1557  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/chromosome="7"  
/map="7q22"  
/cell\_line="HCC-M"  
/cell\_type="hepatocellular carcinoma"  
/tissue\_type="liver"  
1..687  
promoter  
/gene="HCC-1"  
/evidence=experimental  
gene  
1..1557  
/gene="HCC-1"  
305..594  
misc\_feature  
/gene="HCC-1"  
/note="non-functional upstream open reading frame"  
/function="lightens and regulates translation of HCC-1 gene"  
repeat\_region  
446..453  
/rpt\_type=DIRECT  
mrna  
661..1557  
/gene="HCC-1"  
688..1320  
/gene="HCC-1"  
/codon\_start=1  
/product="HCC-1 protein"  
/protein\_id="CAC37950.1"  
/db.xref="GI:13940310"  
/translation="MATEVFLHKIKLAELKOECLARLETKGKIKODLIHLOAYLEE  
HAEEEDVLDSEETETKPIELPKKEEPEPKTVDAEKVYITSEIPTER  
MOKRERNVPSLESKKARARGISVPTKGLSSDNKPMVNDLKEAORGLN  
VSSISSEDEKLEKKRERFGIVTSSAGTGTEDTEAKKKRRERFGIA"  
3'UTR  
1321..1557  
/gene="HCC-1"  
/evidence=experimental  
repeat\_region  
1358..1365  
/rpt\_type=DIRECT  
polyA\_site  
1542  
/gene="HCC-1"  
/evidence=experimental  
BASE COUNT 413 a 365 c 402 g 377 t  
ORIGIN  
Query Match 100.0%; Score 894; DB 9; Length 1557;  
Best Local Similarity 100.0%; Pred. No. 3.3e-201;  
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 1 gggagtgagtgagtgtaacaagatgagcagcagcagcgttgagctccataagctaaag 60  
664 GGGAGTGGAGTGAAGGATGAAGATGGCGACGAGCGGTGAGCTCATTAAGCTAAAG 723  
Qy 61 ctggcgaactaaagcaagaatgtctgtctgtgttggaacccaaggaataagcaa 120  
724 CTTGCCGAACCTAAGCAAGAAATGCTTGCTGCTTTGGACACCAAGGAATTAAGCAA 783  
Db 121 gactttaccagaactccagcagcatatcttgaagaacatgctggaagaaggcaaatgaa 180  
784 GATCTTATCCACAGACTCCAGCATATCTTTGAAGAAACGCTGTAAGAGAGCAAAATGAA 843  
Qy 181 gaagatgacttggaagatgaacacagaagaagaacaaagccattgagctccctgtc 240  
844 GAAGATGTAAGTGGAGATGAAGACAGAGAAAGAAACCAACCCATTGAGCTCCCTGTG 903  
Db 241 aaagaggaagaacccccgaaaaaacctgtgtgtgtgcaagcagaagaagaagtgtgaaa 300  
904 AAAGAGGAAGAACCCCGTGAATAAACTGTGATGTGGCGCAGAGAGAAAGAACTGTGAAA 963  
Qy 301 attacatctgaataaccacagactagagaatgacagaagaaggctgaacgatcaatgta 360  
964 ATTACATCTGAATAATCCACAGACTAGAGAAATGCAGAAAGAGGGCTGAACGATTCATGTA 1023

Qy 361 cctgtgagcttggaagatgaagaagctgtctcggcagcagtaaggtttgggaattcttcaglt 420  
Db 1024 CCTGTGAGCTTGAGAGATGAAGAAAGCTGCTCGGAGAGTAGGTGGATTTCTTCAGT 1083  
Qy 421 ccaacaaaaggtctgtcatctgtataacaaacctatggttaacttggatagctgaagaa 480  
Db 1084 CCACAAAAGGCTGTGATCATCTGATTAACAAACCTATGTTAACTGTGAATGAAGCAA 1143  
Qy 481 agagctcaagaattggtttgtaattgtcttcaactccccaagaagtctgaatgatgag 540  
Db 1144 AAGCTCAAGATTGTGTTGATGATGCTTTCATCTCCAGAAAGTGTGAATGATGAG 1203  
Qy 541 aaactgaaaaaggaaggaagcagatttggattgtcacaaagtcagcttgaacttgaacc 600  
Db 1204 AAACGTAAAGAGAGAGAGGAGGATTTGGATTTGCACAAAGTTGAGCTGAAGCAAC 1263  
Qy 601 acagaggtatacagaagcacaagaagaagaagaagcagagcgtcttggatcttcgtatga 660  
Db 1264 ACAGAGGATACAGAGGCAAAAGAGAGAAAGAGAGAGACCGCTTGGATTCCTGATGA 1323  
Qy 661 aaagttccctgatacttctgtctccagtggtttccattctctctctctctcttgatcac 720  
Db 1324 AAGTTCTGATACCTTCTGTTCTCCAGTGTTTCCATTCTCTCTCTTCTTGTGTCAC 1383  
Qy 721 atataatccaaatgacacagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 780  
Db 1384 AATATATGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1443  
Qy 781 ccccgagtcacatccatgaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 840  
Db 1444 CCCAGGATACATCCATGAACTCGGACGAGTTTACCTTATGTCGTTTACGCTTTAAG 1503  
Qy 841 tctgtgtgttctgttctgtatcatgctgtctgttaataaaaaaataagaaa 894  
Db 1504 TTGTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 1557  
RESULT 2  
AX048092 1071 bp DNA linear PAT 15-DEC-2000  
LOCUS Sequence 86 from Patent WO0070047.  
DEFINITION AX048092  
ACCESSION AX048092.1 GI:11876915  
VERSION  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Yue,H., Tang,Y.T., Lal,P., Reddy,R., Batra,S., Baughn,M.R.,  
Yang,J., Azimzal,Y., Lu,D.A., Au-Young,J. and Shih,L.L.  
TITLE Full-length molecules expressed in human tissues  
JOURNAL Patent: WO 0070047-A 86 23-NOV-2000;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1..1071  
/organism="Homo sapiens"  
/db.xref="taxon:9606"  
/note="Incyte ID No: 4093555CB1"  
BASE COUNT 338 a 190 c 297 g 246 t  
ORIGIN  
Query Match 99.7%; Score 891; DB 6; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 1.7e-200;  
Matches 891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 4 agtgaagtgagtgagtgtaacaagatgagcagcagcagcgttgagctccataagctaaagct 63  
Db 163 ATTGAGTGTAGGGATTAACAAGATGGCGACGACGAGGTGAGCTCATTAAGCTTAAAGCTT 222  
Qy 64 gccgaactaaagcaagaatgtctgtctgtgttggaagcaaggaataagcaagat 123

Db	223	GCCCACTAAAGCAAGATGCTTGTCTGCTGTTGGAGACCAAGGCAATTAACGACAT	282
Qy	124	cttatccacagactccagggacatcttgaagaacatgctgaagaaggagcaatgaagaa	183
Db	283	CTTTATCCACACACATCCAGGCAATCTTGAAGAACATGCTGAAGAGGAGCAAAATGAAGAA	342
Qy	184	gatttactggagatgaacaacaggaagaagaacaacaaagccattgactccctgtcaaa	243
Db	343	GATTTACTGGAGATGTAAMACAGAGGAGAGMAAACAAGCCCATTTGACTCCCTGTCAA	402
Qy	244	gaggaagaacccccctgaaaaaacgttgatgtggtcagcagaagaagaatgggtgaaaaat	303
Db	403	GAGAGAGAACCCCTGAAAAAACTGTTGATGTGACGACGAGAGAAAGTGGTGAATTT	462
Qy	304	acatctgaataacacacagactgagaagaatgcagaagaaggctgaacgattcaatgtacct	363
Db	463	ACATCTGAAATACCAACACACTGAGAGATGAGAGAGAGGCTGAACGATCAATGTACTCT	522
Qy	364	gtgagcttgagagtaagaagaagctgcgcgcgcagctaggtgttgagattcttcacgtcca	423
Db	523	GTGAGCTTGGAGATTAAGAAAGCTGCTCGGACAGCTAGGTTGGGATTTCTTCACTTCCA	582
Qy	424	acaaaaggtctgtcatctgatacaacaacctatggttaacttgataagctgaaggaaaga	483
Db	583	ACAAAAGGCTGTGATCTGATTAACAACCTATGTTAACTTGGATTAAGCTGAAGCAAGAGA	642
Qy	484	gtctaaagaattgtgttgaaatgctcttccaatcccaagaagctgaagaatgagagaa	543
Db	643	GCTCAAAAGATTGTGTTGAATGCTCTTCAATCTCCAGAAAGCTGAAGATGATGAGAAA	702
Qy	544	ctgaaaaaggaagaagcgaattcgggattgtcacaagttcagctggaactggaaccaca	603
Db	703	CTGAAAAAGAGGAAGAGACGATTGGGATTTGTCCACAAGTTCAAGCTGGAACCTGGAACCA	762
Qy	604	gaggaatacagaaggaaagaagaagaagaacagcgtttggatgtgctcgtatgaaaaa	663
Db	763	GAGGATACAGAGGCAAGAGAGAGAGAAAGACAGACGAGCCCTTGGGATTTGCCATGATGAAAA	822
Qy	664	gtctctgatacttctgtctcccaagtgtttcccaattctctccctctctctgtgtacaata	723
Db	823	GTTCCTGATTAATCTTCTGTTCCAGTGTTCACATTTCTTCCTCTCTGTCACATA	882
Qy	724	tatgcctaaatgcacagctcatgtgcctcagctgcgcgaatgaggaacatgtacc	783
Db	883	TATGCTTAATATGCACAGCATGTGCTTACGTCCTCGCAATGAGGAGCATTTACCC	942
Qy	784	caagttacatccatgaactgcgcgcgcagcttgaactatgtctgtttcagctttaagttg	843
Db	943	CAGGTAATTCATGAACTGGGCGACACGCTTTGACTTATTTGCTTTCAGCTTTAAGGTTG	1002
Qy	844	ctgtgtttgtttttgtgtatgtgtgtgttgaataaataaataaataaataa 894	
Db	1003	TTGTGTTTTTTTGTGTTGATTAATGTTGCTTGAATTAATAAATAAATAAAGAAA 1053	
RESULT	3		
LOCUS	BC007099	923 bp	mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, similar to RIKEN CDNA 1110005A23 gene, clone		
ACCESSION	BC007099	MGC:14726 IMAGE:4273903, mRNA, complete cds.	
VERSION	BC007099.1	GI:13937970	
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 923)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (30-APR-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		

REMARK	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
COMMENT	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
	Contact: MGC help desk		
	Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a>		
	Tissue Procurement: CLONTECH		
	CDNA Library Preparation: CLONTECH Laboratories, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web Site: <a href="http://www.shgc.stanford.edu">http://www.shgc.stanford.edu</a>		
	Contact: (Dickson, Mark) <a href="mailto:mcdexall@stanford.edu">mcdexall@stanford.edu</a>		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a>		
Source	Series: IRAL Plate: 21 Row: 9 Column: 18		
	This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.		
	Location/Qualifiers		
	1..923		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="MGC:14726 IMAGE:4273903"		
	/tissue_type="Prostate"		
	/clone_lib="NIH_MGC_83"		
	/lab_host="DH10B"		
	/note="Vector: pDNR-LIB"		
CDS	22..654		
	/codon_start=1		
	/product="Similar to RIKEN CDNA 1110005A23 gene"		
	/protein_id="AA07099.1"		
	/db_xref="GI:13937971"		
	/translation="MATEVELHLKLAELKOECLARGLTEKIKODLIRHLYALEE		
	HAEEANEEDVLGDETEEEKPIELVKEEPEPEKVDAAEKVYKITSIPOTER		
	MOKRAERENVVLSKKKAARAAFGVSSVPTKLSNDKPMVLDLTKRAORFGLN		
	VSSISRSSEDEKREKRRGFIYSSAGVGTMTDTAKRKRAREGRIA"		
BASE COUNT	316 a	156 c	232 g
ORIGIN	219 t		
Query Match	99.5%	Score 889.4;	DB 9; Length 923;
Best Local Similarity	99.9%	Pred. No. 4e-200;	
Matches 890; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	4	agtgagtgagggttaacaagatgagcagcagagcgtgagctccataaagctt	63
Db	1	AGTGAGTGAGGGGTAAACAAGATGGCGACCGAGCGTGGAGCTCCATTAAGCTTAAGCTT	60
Qy	64	gccgaactaaagcaagaatgtctgtctgtgttgtagaaccagaaggaataaagcaagat	123
Db	61	GCCGAACTAAAGCAAGATGCTTCTGCTGTGTTGGAGACCAAGGAATTAAGCAAGAT	120
Qy	124	ctatccacagactccagggatcttgaagaacatgctgaagaaggaggaagaagaaga	183
Db	121	CTTATCCACAGACTCCAGCATATCTTGAAGAACATGCTGAAGAGAGGCAAAATGAAGAA	180
Qy	184	gattgacttgagatgaagaacagaggaagaagaagaagaagccattgagctccctgtcaaa	243
Db	181	GATGTAATGCGGAGATGAAGAAAGAGAGAAACAAAGCCATTGAGACTCCCTGTCAA	240
Qy	244	gaggaagaacccccctgaaaaaaactgtatgtgtgcacagagaagaagaatggtgaaatc	303
Db	241	GAGGAAGAACCCCTGAAAAAACTGTGATGTGCGCAGAGAAAGAGTGGTGAATTT	300
Qy	304	acatctgaataacacacagactgagaagaatgacagaaggagctgaacgattcaatgtacct	363
Db	301	ACATCTGAATTAACACAGCTGAGAGAAATGCAGAAAGAGGCTGAACGATCAATGTACT	360
Qy	364	gtgagcttgagagtaagaagaagctcgcgcgcagctaggttgagattcttcaatgtcca	423

Db	361	GTGAGCTTGAGAGTAGAAGAACCTGCTCGGCGACGTAGGTTGGATTCTTCAGTTCCA	420
Qy	424	acaaaagctctctcatctgatacaaaacctatggttaacttgataagctgaaggaaga	483
Db	421	ACAAAAGGCTCTCATCTGATACAAACCTATGTTAACTTGATAGCTGAAGGAAGA	480
Qy	484	gtcacaagaattggttgaatgctctctcaatctccacagaagctcgaagatagagaa	543
Db	481	GCTCAAGATTGGTTGATGATCTCTTCATCTCCAGAAACTCGAAGATGATGAGAA	540
Qy	544	ctgaaaagaaggaagagcgaatttgagatgtcacaaagtccgagcctggaactgga	603
Db	541	CTGAAAAGAGGAGAGAGGATTGGGATTGTTCACAAAGTTGACGTGGAACCTGGA	600
Qy	604	gagagatagaggaagaaagagaagagagagcgccttggagatgcccgtatgaaa	663
Db	601	GAGGATACAGAGGCANAAGAGAGAAAGACAGACGCTTGGGATGCTCGATGAGA	660
Qy	664	gttcctgatactctgtctctcgaatgtttccattctctctctctctctcttgta	723
Db	661	GTTCTGATACTTTCTGTCTCCTCAGGTTCATTTCTCCTCTCTTGTGTCACATA	720
Qy	724	tatgctcaaatgacagatcatgtgctctcagctcctcgcgaatgagagacatgccc	783
Db	721	TATGCTTAATGTCACAGTCATGTGCTGCTGCTGCTGCAATGAGGAGCATGTACC	780
Qy	784	caggtacatcatcagatcgcgcagcagcttgactaatgtcgtttcagcttcaagttg	843
Db	781	CAGGTCATCATGACATGCGGACGAGTTTACCTTATTGCTGTTTACCTTTAAGTTG	840
Qy	844	tttgttttttgttttgatattgctgtgttaataaaaaaataagaa 894	
Db	841	TGTGTTTGTGTTTGTGATTATGTTGCTGTTAATAAAAAAATAAGAAA 891	
RESULT	4		
LOCUS	BC016941	1818 bp	mRNA
DEFINITION	Homo sapiens, clone MGC:21452 IMAGE:3448446, mRNA, complete cds.		
ACCESSION	BC016941		
VERSION	BC016941.1	GI:16877383	
KEYWORDS	MGC.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 <bases 1 to 1818>		
JOURNAL	Strausberg, R.		
REMARK	Direct Submission		
COMMENT	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
	Contact: MGC help desk		
	Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a>		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>		
	Contact: (Dickson, Mark) <a href="mailto:mdickpaxil.stanford.edu">mdickpaxil.stanford.edu</a>		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Series: IRAK Plate: 20 Row: a Column: 13		
	This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not Identity to protein.		
FEATURES	Location/Qualifiers		

source	1. .1818		
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="MGC:21452 IMAGE:3448446"			
/tissue_type="Placenta, choriocarcinoma"			
/clone_1id="NIH_MGC_10"			
/lab_host="DH10B"			
/note="Vector: pCMV-SPORT6"			
401. .1570			
/codon_start=1			
/product="unknown (protein for MGC:21452)"			
/protein_1d="AAH16941.1"			
/db_xref="GI:16877384"			
/translation="MKRAENELSRVNEFLSKLQDDLKEMANTMCSRCGRHREF MDREPSARVACENRLHAEEDGPMFESSMLGKITFYFALMDKRYDITPMACORV GISPDTHRVPHYISRGSRIPGTRGROTPADPADLQDFISRIFOVPGQMPGNGF AADPPAPDAANAASPNSTVYPGEKPKPRKKRLALAKOBLARGLFETGIRKODLIHRQ AYLEHAAEENAEEDVLGDETEEBEETPIELPVAEEPEEPTVDVAEKVKVLTSEI PTERMOKRAERFNVPSVLESKKVAAARAFGISVPTPKGISDNRKPMVNLDKLERAQ RFGLNVSISRSSEDEKLRKKRERFCIVTSSAGCTGTEDTEAKRRRAERFCTA"			
BASE COUNT	532 a	373 c	494 g 419 t
ORIGIN			
Query Match	93.3%;	Score 833.8;	DB 9; Length 1818;
Best Local Similarity	99.8%;	Pred. No. 5.7e-187;	
Matches 835; Conservative	0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	58	aagcttgcggaactaaagaagaatgcttgcctgtgtgttggaaaccaaggaataag	117
Db	971	AAACTTGGCGGAAGCTAAAGCAAAATGCTGCTGTGTTGGAAACCAAGGAAATAAG	1030
Qy	118	caagatcttatccagagactccagcagcatcttgaagaacatgctggaagagagcaaat	177
Db	1031	CAGATCTTATCCACAGACTCCAGCATATCTTGAAGAACAATGCTGAAGAGAGCAAT	1090
Qy	178	gaagaagatgactggyagatgaaacagagagaagaagaacaagcccatgagctccct	237
Db	1091	GAAAGAGATGACTGAGATGATAAACAAGAGAAACAAGCCCATTTAGCTCCCT	1150
Qy	238	gtcaaaagagaagaaacccctctgaaaaactgtgtgtgtgtgtgtgtgtgtgtgtgt	297
Db	1151	GTCAAAGAGAGAACCCCTGAAAAACTGTGATGTGGCAGCAGAGAGAAATGTGGG	1210
Qy	298	aaaattacatctgaaataccacagactgagagaaatgcagaagaagctgaagatcaat	357
Db	1211	AAAATTACTCTGAAATATACACAGACTGAGAAATGCGAGAAAGGCTGAAGATTCAAT	1270
Qy	358	gtacctgtgagcttggagataagaagctgctcggaagctaggttttggattctctca	417
Db	1271	GTAAGTGTGAGCTTGAGAGTAAAGAGTTCGCGGACGTAGGTTGGGATTTCTTCA	1330
Qy	418	gttccaacaaagctctgcatctgatacaaacctatggttaacttgataagctgaag	477
Db	1331	GTTCCAAACAAAGCTCTGATCTGATTAACAACCTATGCTTGAATGAGTGAAG	1390
Qy	478	gaagaagctcaagaatttggtttgaatgctcttcaatctccagaagaagctgaagatg	537
Db	1391	GAAAGAGCTCAAAAGTTTGGTTTGAATGCTCTTCATCTCCAGAAAGCTGAACATGAT	1450
Qy	538	gagaactgtaaaagaagaagagcgaatttggatgtgtacaaagtccagctggaactgga	597
Db	1451	GAGAAACTGAAGAGAGAGAGGAGGATTGGATTGTCACAAGTTGACGCTGAACCTGA	1510
Qy	598	accacagagatagagaggaagaaagagaaagcagagcgccttggagattgtccgga	657
Db	1511	ACCACAGAGATACAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1570
Qy	658	tgaaagttccgtgaactctgtctcagtggtttccattccctctctctctctgtgt	717
Db	1571	TGAAAAGTTCCGTGATATCTTCTGTTCTCCAGTGTTTTCATTTCTCTCTTCTTGT	1630
Qy	718	caatataatgcttaaatgacagatcatgtgtcctcgtcctcgcaatgaaggagcat	777

Db	1631	CACATATAGCCTTAATGACACAGTCATGTGGCTAGTCTCGTCTGCATAGGAGCAT	1690
Oy	778	gtaccacaggtacatcatgaactgcgcaagcagittgactattgtcttcaagtta	837
Db	1691	GTAACCACAGGATACATCACTGAACCTCCGCCAGAGTTTGACTATTGCTGTTTACACTTTA	1750
Oy	838	agggttgctgttttgttttgttatgatcgttgctgtttaataaaaaaaatagaana	894
Db	1751	AGGTTGTTGTTGTTTTTTGTTTGTATTATGTTCTGTTTAATAAAAAAAATAGAANA	1807
RESULT	5	.	
LOCUS	AC073063	84113 bp DNA linear PRI 09-JAN-2002	
DEFINITION	Homo sapiens BAC Rpl1-136B3 from 7, complete sequence.		
ACCESSION	AC073063		
VERSION	AC073063.12	GI:15638765	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 84113)		
JOURNAL	Sulston,J.E. and Waterston,R.		
AUTHORS	Toward a complete human genome sequence		
MEDLINE	Genome Res. 8 (11), 1097-1108 (1998)		
REFERENCE	99063792		
AUTHORS	2 (bases 1 to 84113)		
TITLE	Armstrong,J., Maupin,R. and Long,N.		
JOURNAL	The sequence of Homo sapiens BAC clone Rpl1-136B3		
AUTHORS	Unpublished (2001)		
TITLE	3 (bases 1 to 84113)		
JOURNAL	Waterston,R.H.		
AUTHORS	Direct Submission		
TITLE	Submitted (08-JUN-2000) Genome Sequencing Center, Washington		
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
AUTHORS	MO 63108, USA		
TITLE	4 (bases 1 to 84113)		
JOURNAL	Waterston,R.H.		
AUTHORS	Direct Submission		
TITLE	Submitted (18-SEP-2001) Genome Sequencing Center, Washington		
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis,		
AUTHORS	MO 63108, USA		
TITLE	5 (bases 1 to 84113)		
JOURNAL	Waterston,R.		
AUTHORS	Direct Submission		
TITLE	Submitted (09-JAN-2002) Department of Genetics, Washington		
JOURNAL	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA		
AUTHORS	On Sep 18, 2001 this sequence version replaced g1:113431102.		
TITLE	----- Genome Center		
JOURNAL	Center: Washington University Genome Sequencing Center		
AUTHORS	Center code: WUGSC		
TITLE	Web site: http://genome.wustl.edu/gsc		
JOURNAL	Contact: saplens@watson.wustl.edu		
AUTHORS	----- Summary Statistics		
TITLE	Center project name: H_NH0136B03		
JOURNAL	-----		

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NIGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7> , send <mailto:egreen@nhgri.nih.gov> , or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Fritngen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://pacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the right is RP11-114P12, 2000 bp overlap; the clone sequenced to the left is GS1-259H13, 200 bp overlap. Actual start of this clone is at base position 96082 of GS1-259H13; actual end is at base position 84113 of RP11-136B3.

FEATURES  
source  
1. .84113  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7"  
/clone="RP11-136B3"  
/clone\_id="RPCI-11"  
22. .56  
/rpt\_family="AT\_rich"  
46. .330  
/rpt\_family="Alu"  
335. .405  
/rpt\_family="MER1\_type"  
545. .613  
/rpt\_family="MER1\_type"  
614. .780  
/rpt\_family="Alu"  
781. .1078  
/rpt\_family="Alu"  
1080. .1114  
/rpt\_family="MER1\_type"  
1115. .1406  
/rpt\_family="Alu"  
1407. .1540  
/rpt\_family="MER1\_type"  
1833. .2211  
/note="match to EST BE147565 (NID:g8610289)"  
2322. .2394  
/note="match to EST BE147565 (NID:g8610289)"  
2337. .2489  
/rpt\_family="Alu"  
2497. .2619  
/rpt\_family="MIR"  
2620. .2644  
/rpt\_family="(T)n"  
2926. .2991  
/rpt\_family="MIR"  
3035. .3355  
/rpt\_family="Alu"  
3374. .3395  
/rpt\_family="(TTTTRA)n"  
3552. .3852  
/rpt\_family="Alu"  
3853. .3972  
/rpt\_family="Alu"  
4009. .4388  
/rpt\_family="T2\_type"  
4419. .4731

```
repeat_region /rpt_family="Alu" 4786..5085
rpt_family="Alu"
repeat_region /rpt_family="Alu" 5203..5238
rpt_family="L1"
repeat_region /rpt_family="L1" 5239..5651
rpt_family="L1"
repeat_region /rpt_family="L1" 5642..5663
rpt_family="TAAAn"
misc_feature /note="match to EST BE795688 (NID:g10216886)" 5885..6027
misc_feature /note="match to EST BG722929 (NID:g14002116)" 5887..6025
misc_feature /note="match to EST BG722929 (NID:g14002116)" 5887..6025
misc_feature /note="similar to EST BG917776 (NID:g14298252)" 5990..6025
misc_feature /note="match to EST BG748931 (NID:g14059584)" 6106..6382
repeat_region /rpt_family="L2" 6330..6334
misc_feature /note="match to EST BF894562 (NID:g12286021)" 6383..6682
repeat_region /rpt_family="Alu" 6683..6744
repeat_region /rpt_family="L2" 6745..7046
repeat_region /rpt_family="Alu" 7017..7042
repeat_region /rpt_family="(GAA)n" 7047..7139
misc_feature /rpt_family="L2" 7251..7849
repeat_region /note="match to EST AV717144 (NID:g10814296)" 7435..7732
repeat_region /rpt_family="Alu" 7735..8052
repeat_region /rpt_family="Alu" 7907..8001
misc_feature /note="match to EST AV717144 (NID:g10814296)" 7927..8032
misc_feature /note="match to EST BE147565 (NID:g8610289)" 8017..8086
repeat_region /rpt_family="A-rich" 8118..8234
repeat_region /rpt_family="Alu" 8219..8241
repeat_region /rpt_family="A-rich" 8245..8543
repeat_region /rpt_family="Alu" 8300..8302
misc_feature /note="match to EST AV717144 (NID:g10814296)" 8713..9008
repeat_region /rpt_family="Alu" 9055..9190
repeat_region /rpt_family="Alu" 9191..9501
repeat_region /rpt_family="Alu" 9472..9504
repeat_region /rpt_family="A-rich" 9502..9674
repeat_region /rpt_family="Alu" 9648..9686
repeat_region /rpt_family="(GAA)n" 9714..10005
repeat_region /rpt_family="Alu" 10083..10168
repeat_region /rpt_family="L1" 10171..10459
```

QY	12	gaggggttaacaaagatgaggagaccggagagcggtggagcttcataagctaaagctggccgaact	71
Db	74969	GAGGGGTAAACAAGATGGTAACTGAAATGATGTGGAGCTCCATTAAGCTGAAGCTTGCTAAACT	75028
QY	72	aaagcaagaatgctctgtgtcgtgtgtgttgagaccagaaggaaataagcaagaatctctatcca	131
Db	75029	AAAGCAAGAATGTCTT-CTTGTGTGTGGAGCCAAAGGGAATPAAAGCAAGATCTTATCCA	75087
QY	132	caagctccaagcagatctcttgaaagaacatgcttgaagaggccaatgaaagaatgtaact	191
Db	75088	CAGACTCCAGGCAATATCTTGAAGAAACATGCTCAAGAGAGGCAAAAT---GAAGATGTAAT	75144
QY	192	gggagatgaacaacagaggagaaagaacaagcccatgagctccctgtccaaggagaa	251
Db	75145	AGGAAGTGAACACAGAAAGAAAAA-----AACCCCTTGCAATATGGAAGA	75191
QY	252	accacctg-aaaaaacgttgtatgtgagcagcagagaaagaatgggtgaanaatlacatctg	310
Db	75192	ACCCCTGTAAGAAAAAAGTGTGATGTGGCAGCAGAGAAAGAAAGTGATGAATTAATCTTG	75251
QY	311	aaatccacacagactgagaggaatgcaagaagaggctgaaacgattcaatgtacctgtgagct	370
Db	75252	AAATCCACAGGCTTAGAGAAATGCAAGAAAGGGCCGAATTCAGTGTACTCTGTGACT	75311
QY	371	tggagaatgaagaagactgctcgagcagctaaagtgtggagatctctcaatctcacaacaaag	430
Db	75312	TGGAGATGAGAAACTCTCTCAGGACGATAGGATTGGGATTTCTTCACTTCCAAACAAA-	75370
QY	431	gtctgtcatctgataacaacctatglttaacttgataaagctgaaggaaagagctcaaa	490
Db	75371	GTCTGTCACTCTG--AACACACTATAGGTTAACTTGTGATTAACCCGAAGGAAAGACTCTAAA	75428
QY	491	gatttggatttgaatgtctcttcaatctccagaagaatctgaagatgatatgaagaactgaaa	550
Db	75429	GATTTGGTTTGAATGTCTCTTCAATTTCCAGAAACTCTGAAGATGATATAGAAACTGAAAA	75488
QY	551	agaaggaagagcgatcttggagatgtgacaagtctcaactgaaactgaaaccaagaagata	610
Db	75489	AGAGGAAGAGCGATTGGGATTGTCAACAAGTTCAGCTGGAACCTGGAACCCACAGAGATA	75548
QY	611	cagagcgaaagaagaggaagaagaagagcgcttggagatggcgcgatgaaagttctgt	670
Db	75549	CAGAGCGAAAGAGAGAAAAAGAGAGAGCGCTTGGCATTTGCCGATGAAGAGGTTCTTG	75608
QY	671	ataacttctgtctccagtggttctccattctctctctctctctctgtgtacatatagct	730
Db	75609	ATACCTTCTGTTCTCCAGGTGTTTCCACTTCTCTCATTTCTTGTGTGTCACATATATATA	75668
QY	731	aaatgcacagtcatagtgcctacgctccgtccctgcgaatgaaggagacatgtaccceagttac	790
Db	75669	AA-----TGACGTCTTGCGCTTGTGCAATAGAGAGAGCAGGTACCCACAGGTAA	75712
QY	791	atccatgaactgcggcagcagagttgaactatgtgcgttccagctttaaagtgctgtgtgt	850
Db	75713	AACGTGAACCTCCCGCGACGATTGTACTTATTTGCTGTTTCAAGCTTTAAAGGTGTGTGT	75772
QY	851	tttgttttgataatgtgtgtgt 875	
Db	75773	TTTGTGTTTGAATATATGCTGTT 75797	

Query Match	72.48;	Score 647.4;	DB 9;	length 84113;
Best Local Similarity	90.48;	Pred. No. 6.9e-143;		
Matches 782;	Conservative 0;	Mismatches 46;	Indels 37;	Gaps 7;

AC022200 148152 bp DNA linear HTG-09-MAY-2001  
Homo sapiens chromosome 8 clone RP11-262017 map 8, WORKING DRAFT  
SEQUENCE, 4 unordered pieces.  
AC022200  
AC022200.6 GI:13518207  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE AUTHORS 1 (bases 1 to 148152)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 8, clone RP11-262017  
JOURNAL TITLE Unpublished  
REFERENCE 2 (bases 1 to 148152)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckert, R., Beda, F., Boguslavsky, L., Borkhagalter, B., Brown, A., Burkett, G., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeRitellano, K., Dewar, K., Domino, M., Doyle, M., Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Hearford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kam, L., Karatas, A., Klein, J., Landers, T., Lenoczky, J., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McQuirk, A., McKernan, K., McHeaters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, D., Testaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

JOURNAL TITLE Direct Submission  
COMMENT Submitted (26-Jan-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 3, 2001 this sequence version replaced g1:13488035.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L5196  
Center clone name: 262\_O\_17  
----- Summary Statistics  
Sequencing vector: M13; M77815: 63% of reads  
Sequencing vector: Plasmid; n/a: 37% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 146959 bases at least Q40  
Consensus quality: 147487 bases at least Q30  
Consensus quality: 147658 bases at least Q20  
Insert size: 145000; agarose-fp  
Insert size: 147852; sum-of-contigs  
Quality coverage: 8.8 in Q20 bases; agarose-fp  
Quality coverage: 8.7 in Q20 b.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 40424: contig of 40424 bp in length  
\* 40425 40524: gap of 100 bp  
\* 40525 48708: contig of 8184 bp in length  
\* 48709 48808: gap of 100 bp  
\* 48809 73107: contig of 24299 bp in length  
\* 73108 73207: gap of 100 bp  
\* 73208 148152: contig of 74945 bp in length.  
\* Location/Qualifiers  
1. 148152  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-262017"  
/clone\_lib="RPct-11 Human Male BAC"  
1. 40424  
misc\_feature

[illegible]

AP001207 153936 bp DNA linear PRI 24-JUN-2000  
 LOCUS Homo sapiens genomic DNA, chromosome 8q23, clone:KB1562D12.  
 ACCESSION AP001207  
 VERSION AP001207.3 GI:8698839  
 KEYWORDS  
 SOURCE Homo sapiens pre-pro-B cell cell\_line:FLB 14 - 14 DNA,  
 clone\_lib:Keio BAC library clone:KB1562D12.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 153936)  
 AUTHORS Shimizu,N and Asakawa,S.  
 JOURNAL Homo sapiens chromosome clone BAC KB1562D12 on 8q23  
 REFERENCE 2 (bases 1 to 153936)  
 AUTHORS Shimizu,N. and Asakawa,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-FEB-2000) to the DDBJ/EMBL/GenBank databases.  
 Nobuyoshi Shimizu, Keio University, School of Medicine, Molecular  
 Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan  
 (E-mail:nshimizu@med.keio.ac.jp, Tel:81-3-3351-2370,  
 Fax:81-3-3351-2370)  
 COMMENT On Jun 24, 2000 this sequence version replaced gi:8096488.  
 FEATURES  
 source location/Qualifiers  
 1..153936  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="FLB 14 - 14"  
 /cell\_type="pre-pro-B cell"  
 /chromosome="8"  
 /clone="KB1562D12"  
 /clone\_lib="Keio BAC library"  
 /map="8q23"  
 9..330  
 /rpt\_family="LTR33"  
 /evidence-not\_experimental  
 repeat\_region 2020..2312  
 /rpt\_family="AluJb"  
 /evidence-not\_experimental  
 repeat\_region 2403..2462  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 2795..2827  
 /rpt\_family="AT-rich"  
 /evidence-not\_experimental  
 repeat\_region 3134..3161  
 /rpt\_family="AT-rich"  
 /evidence-not\_experimental  
 repeat\_region 3467..3726  
 /rpt\_family="MIR"  
 /evidence-not\_experimental  
 repeat\_region 5475..5696  
 /rpt\_family="MER58A"  
 /evidence-not\_experimental  
 repeat\_region 6083..6111  
 /rpt\_family="AT-rich"  
 /evidence-not\_experimental  
 repeat\_region 6827..7002  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 7203..7308  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 7433..7546  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 8592..8715  
 /rpt\_family="MIR"  
 /evidence-not\_experimental

repeat\_region 8719..8757  
 /rpt\_family="AT-rich"  
 /evidence-not\_experimental  
 repeat\_region 8770..9035  
 /rpt\_family="AluJo"  
 /evidence-not\_experimental  
 repeat\_region complement(10285..10480)  
 /rpt\_family="MER3"  
 /evidence-not\_experimental  
 repeat\_region complement(11532..11678)  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 11906..12004  
 /rpt\_family="(TAA)n"  
 /evidence-not\_experimental  
 repeat\_region complement(12010..12084)  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region complement(12089..12156)  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region 12747..13156  
 /rpt\_family="MTL12"  
 /evidence-not\_experimental  
 repeat\_region 13258..13290  
 /rpt\_family="AT-rich"  
 /evidence-not\_experimental  
 repeat\_region 13337..13502  
 /rpt\_family="MIR"  
 /evidence-not\_experimental  
 repeat\_region 13537..13643  
 /rpt\_family="L2"  
 /evidence-not\_experimental  
 repeat\_region complement(13709..14015)  
 /rpt\_family="AluJo"  
 /evidence-not\_experimental  
 repeat\_region 14796..14829  
 /rpt\_family="(A)n"  
 /evidence-not\_experimental  
 repeat\_region 15290..15526  
 /rpt\_family="MIR"  
 /evidence-not\_experimental  
 repeat\_region 15834..15998  
 /rpt\_family="LIME1"  
 /evidence-not\_experimental  
 repeat\_region 17379..18211  
 /rpt\_family="LIME1"  
 /evidence-not\_experimental  
 repeat\_region 18212..18333  
 /rpt\_family="AluX"  
 /evidence-not\_experimental  
 repeat\_region 18334..18362  
 /rpt\_family="(CA)n"  
 /evidence-not\_experimental  
 repeat\_region 18363..18517  
 /rpt\_family="AluX"  
 /evidence-not\_experimental  
 repeat\_region 18518..18543  
 /rpt\_family="(CAA)n"  
 /evidence-not\_experimental  
 repeat\_region 18544..18814  
 /rpt\_family="LIME1"  
 /evidence-not\_experimental  
 repeat\_region complement(18991..19309)  
 /rpt\_family="AluSg"  
 /evidence-not\_experimental  
 repeat\_region 20149..20221  
 /rpt\_family="L1MC5"  
 /evidence-not\_experimental  
 repeat\_region 20375..20669  
 /rpt\_family="AluY"  
 /evidence-not\_experimental  
 repeat\_region 20673..21016



	/rpt_family="L1MC5"	/evidence-not_experimental
repeat_region	21017..21314	/rpt_family="Alusx"
	/evidence-not_experimental	21315..21613
repeat_region	/rpt_family="L1MC5"	/evidence-not_experimental
repeat_region	22283..22576	/rpt_family="Alusp"
	/evidence-not_experimental	24614..24918
repeat_region	/rpt_family="Alusc"	/evidence-not_experimental
repeat_region	26438..26697	/evidence-not_experimental
	/rpt_family="Aluy"	/evidence-not_experimental
repeat_region	26893..26919	/rpt_family="(TTTC)n"
	/evidence-not_experimental	27106..27295
repeat_region	/rpt_family="MIR"	/evidence-not_experimental
repeat_region	27296..27337	/rpt_family="AT_rich"
	/evidence-not_experimental	27464..27948
repeat_region	/rpt_family="L1MB3"	/evidence-not_experimental
repeat_region	27994..28335	/rpt_family="L1PA16"
	/evidence-not_experimental	28338..28360
repeat_region	/rpt_family="AT_rich"	/evidence-not_experimental
repeat_region	28364..28398	/rpt_family="L1PA7"
	/evidence-not_experimental	28400..28447
repeat_region	/rpt_family="A-rich"	/evidence-not_experimental
repeat_region	29033..29233	/rpt_family="L2"
	/evidence-not_experimental	29386..29566
repeat_region	/rpt_family="L2"	/evidence-not_experimental
repeat_region	30360..30407	/rpt_family="(TGGA)n"
	/evidence-not_experimental	31826..32126
repeat_region	/rpt_family="AlubB"	/evidence-not_experimental
repeat_region	32360..32668	/rpt_family="AlucC"
	/evidence-not_experimental	
Query Match	61.5%:. Score 549.6; DB 9;	Length 153936;
Best Local Similarity	86.8%; Pred No. 9.4e-120;	
Matches 665; Conservative 0; Mismatches 89; Indels 12; Gaps 5;		
DB 132	cagactccaggcatatcttgaagaacatgctgaagaggaagcacaataagaagaagtgtact	191
DB 49353	CTGGGTCACAGGCATATTTCAGGGCCATGCGTAACAGG-----CAAAATGAAGAATGTACT	49406
DB 192	gggagatgaaacagagaagaagaacaag-cccatgaagtcaccttcaagaagaaag	250
DB 49407	GGGAGATGAACAACAGAGGAAGAAACAAGGCCCATGAATCCCTGCAAAAGAGGAAG	49466
DB 251	aacccttgaaaaaaccttgtatgtggagcagagaagaagaagtgtgaaaatatcatctg	310
DB 49467	AACCTCCCTGAATAAACTTTGACGTGGAGCGAGAGCAAGAACTGGTAAAAAATTACATCTG	49526
DB 311	aatatcacacagactlgaagaaatgcagaagagggctgaaccgattcaatgtlaacttgaagct	370

Db	Accession	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS
Db	49527	AAATACCCAGGAGAGAGAAATGCAGAAAGAGGCGCTGAACCAAGTTCATTTACCCGTGAGCT	166471 bp	DNA	linear	HTG 25-MAR-2001							
Qy	371	tgagagaagaagaagctctcgcgcagcctaagtttggatttcttcagttccacaagaag	166471 bp	Homo sapiens chromosome 8 clone RP11-373L22 map 8, WORKING DRAFT	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49587	TGTAAAGTAAAGAAACTCCCGGAGCAGTACGTTTGGACCTTTCACATTTCAACAAAG	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	431	gtctgtcatctgtataacaacacctatglttaacttggataaagctgaaggaagagctcaaa	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49647	GTCCTGCTCATCTTAACAAACCAACTATCATGAAGACCTGATAGCTTAAGGAAGAGAGCTCAAA	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	491	gatttggatttgaatgtctcttcaatcttcagaagttcgaagatgtatgaagaactgaa	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49707	GATTTGGTTTGAATGGCTCTTCACCTCCAGAAATCTGAAAGATATATAGAAACGTA	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	551	agaagaagaagcagatttggatttgcacaaagtctacgcctggaacttga--accacagaag	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49767	GGAGGAAGAGACTGATTTGGGATTATCACTGGTTACGCTGGAAACGAAAGAAACACAGAGG	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	608	atacagaagcaagaagaagaagaagacagagcgcttggatttgcctgtatgaagaaagttc	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49827	ATACAGAGAGCAAAAGAAATAGAAAGAGCGCTGCGCTATGGGATCTCGTGAAGAAAGTTTC	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	668	ctgaacttctgtcttctccagtggtttccatttctctcctcttcttcttggtaacataatg	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49887	ATGCGCTTTCTGTTCTCTCAGTGGTTTCCATTCTGTAATCTTCGCGTCACAATATATA	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	728	cctaactcacagtcacgtctgcctacgtctcgtcgcgaatgagggagatgaacccaag	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	49947	CCTAAATGCACAGTACGTACTACTTACCTCTGCTTGCATGAGGCGACATATGTTC	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	788	tacalcctgaactgcgcgcagcagcttgcactatctgcctgttcaagcttlaagttgtgt	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	50007	TACATGTGGAAACCTCCAGACGAAATTTGACTATTTGATGTTGCCAAC--TTAAAGCTGTGT	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Qy	848	gttttgttttgaattatgttgccttgccttaataaaaaataagaag	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				
Db	50066	GTTTTGTTTTTGATT--TTTTCCTTGTATGTTAAAAA	166471 bp	Homo sapiens chromosome 8, clone RP11-373L22	AC091052	AC091052.1	GI:13446282	HTG: HTGS_PHASE1; HTGS_DRAFT.	human.				



REFERENCE Mammalia: Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 149312)  
 AUTHORS Waterston, R.H.  
 TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 149312)  
 . AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-JUN-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

## COMMENT

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 ----- Project Information -----  
 Center project name: H\_NH0373L22  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 100%  
 Sequencing vector: plasmid; 0%  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 0% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 131361 bases at least Q40  
 Consensus quality: 136097 bases at least Q30  
 Consensus quality: 138572 bases at least Q20  
 Insert size: 14500; agarose-fp  
 Insert size: 145612; sum-of-contigs  
 Quality coverage: 3.09 in Q20 bases; agarose-fp  
 Quality coverage: 3.17 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 38 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1352: contig of 1352 bp in length  
 \* 1353 1452: gap of unknown length  
 \* 1453 3558: contig of 2106 bp in length  
 \* 3559 3658: gap of unknown length  
 \* 3659 5237: contig of 1579 bp in length  
 \* 5238 5337: gap of unknown length  
 \* 5338 6735: contig of 1398 bp in length  
 \* 6736 6835: gap of unknown length  
 \* 6836 8351: contig of 1516 bp in length  
 \* 8352 8451: gap of unknown length  
 \* 8452 9956: contig of 1505 bp in length  
 \* 9957 10056: gap of unknown length  
 \* 10057 11502: contig of 1446 bp in length  
 \* 11503 11602: gap of unknown length  
 \* 11603 12856: contig of 1254 bp in length  
 \* 12857 12956: gap of unknown length  
 \* 12957 14103: contig of 1147 bp in length  
 \* 14104 14203: gap of unknown length  
 \* 14204 15942: contig of 1739 bp in length  
 \* 15943 16042: gap of unknown length  
 \* 16043 17589: contig of 1547 bp in length  
 \* 17590 17689: gap of unknown length  
 \* 17690 19679: contig of 1990 bp in length  
 \* 19680 19779: gap of unknown length  
 \* 19780 21444: contig of 1665 bp in length  
 \* 21445 21544: gap of unknown length  
 \* 21545 23225: contig of 1681 bp in length  
 \* 23226 23325: gap of unknown length  
 \* 23326 25332: contig of 2007 bp in length  
 \* 25333 25432: gap of unknown length  
 \* 25433 27188: contig of 1757 bp in length  
 \* 27189 27289: gap of unknown length  
 \* 27290 29855: contig of 2566 bp in length

FEATURES	source
misc_feature	29856 29955: gap of unknown length
misc_feature	29956 32915: contig of 2960 bp in length
misc_feature	32916 33015: gap of unknown length
misc_feature	33016 35495: contig of 2480 bp in length
misc_feature	35496 35595: gap of unknown length
misc_feature	35596 38408: contig of 2813 bp in length
misc_feature	38409 38508: gap of unknown length
misc_feature	38509 41815: contig of 3307 bp in length
misc_feature	41816 41915: gap of unknown length
misc_feature	41916 45472: contig of 3557 bp in length
misc_feature	45473 45572: gap of unknown length
misc_feature	45573 49293: contig of 3721 bp in length
misc_feature	49294 49393: gap of unknown length
misc_feature	49394 52896: contig of 3503 bp in length
misc_feature	52897 52996: gap of unknown length
misc_feature	52997 56786: contig of 3790 bp in length
misc_feature	56787 56886: gap of unknown length
misc_feature	56887 62170: contig of 5284 bp in length
misc_feature	62171 62270: gap of unknown length
misc_feature	62271 68265: contig of 5995 bp in length
misc_feature	68266 68365: gap of unknown length
misc_feature	68366 72866: contig of 4501 bp in length
misc_feature	72867 72966: gap of unknown length
misc_feature	72967 77441: contig of 4475 bp in length
misc_feature	77442 82838: contig of 5297 bp in length
misc_feature	82839 82938: gap of unknown length
misc_feature	82939 87223: contig of 4285 bp in length
misc_feature	87224 87323: gap of unknown length
misc_feature	87324 93316: contig of 5993 bp in length
misc_feature	93317 93417 98918: contig of 5502 bp in length
misc_feature	98919 99018: gap of unknown length
misc_feature	99019 105925: contig of 6907 bp in length
misc_feature	105926 106025: gap of unknown length
misc_feature	106026 114510: contig of 8485 bp in length
misc_feature	114511 114610: gap of unknown length
misc_feature	114611 124690: contig of 10080 bp in length
misc_feature	124691 124790: gap of unknown length
misc_feature	124791 136488: contig of 11699 bp in length
misc_feature	136489 136589: gap of unknown length
misc_feature	136590 149312: contig of 12723 bp in length.
FEATURES	Location/Qualifiers
source	1. 149312
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="8"
	/clone="RP11-373L22"
	1. 1352
misc_feature	/note="assembly_name:Contig23"
misc_feature	/note="assembly_name:Contig25"
misc_feature	3659. 5237
misc_feature	/note="assembly_name:Contig26"
misc_feature	5338. 6735
misc_feature	/note="assembly_name:Contig27"
misc_feature	6836. 8351
misc_feature	/note="assembly_name:Contig28"
misc_feature	8452. 9956
misc_feature	/note="assembly_name:Contig30"
misc_feature	10057. 11502
misc_feature	/note="assembly_name:Contig31"
misc_feature	11603. 12856
misc_feature	/note="assembly_name:Contig32"
misc_feature	12957. 14103
misc_feature	/note="assembly_name:Contig34"
misc_feature	14204. 15942
misc_feature	/note="assembly_name:Contig35"
misc_feature	16043. 17589
misc_feature	/note="assembly_name:Contig36"
misc_feature	17690. 19679
misc_feature	/note="assembly_name:Contig37"
misc_feature	19780. 21444

	misc_feature	/note="assembly_name:Contig38" 21545..23325	
	misc_feature	/note="assembly_name:Contig39" 23326..25332	
	misc_feature	/note="assembly_name:Contig40" 25433..27189	
	misc_feature	/note="assembly_name:Contig41" 27290..29855	
	misc_feature	/note="assembly_name:Contig42" 29956..32915	
	misc_feature	/note="assembly_name:Contig43" 33016..35495	
	misc_feature	/note="assembly_name:Contig44" 35596..38408	
	misc_feature	/note="assembly_name:Contig45" 38509..41815	
	misc_feature	/note="assembly_name:Contig46" 41916..45472	
	misc_feature	/note="assembly_name:Contig47" 45573..49293	
	misc_feature	/note="assembly_name:Contig48" 49394..52896	
	misc_feature	/note="assembly_name:Contig49" 52997..56786	
	misc_feature	/note="assembly_name:Contig50 clone_end:sp6 vector_end:right" 56887..62170	
	misc_feature	/note="assembly_name:Contig51" 62271..68265	
	misc_feature	/note="assembly_name:Contig52" 68366..72866	
	misc_feature	/note="assembly_name:Contig53" 72967..77441	
	misc_feature	/note="assembly_name:Contig54" 77542..82838	
	misc_feature	/note="assembly_name:Contig55" 82939..87223	
	misc_feature	/note="assembly_name:Contig56" 87324..93316	
	misc_feature	/note="assembly_name:Contig57" 93417..98918	
	misc_feature	/note="assembly_name:Contig58" 99019..105925	
	misc_feature	/note="assembly_name:Contig59"	
Query Match	61.4%;	Score 548.6;	DB 2: Length 149312;
Best Local Similarity	86.8%;	Pred. No. 1.6e-119;	
Matches 664;	Conservative 0;	Mismatches 89;	Indels 12; Gaps 5
QY	132	cagactccagcgatatactcttgaagaacatgctgaagagagcgcaaatgaagaagatgact	191
Db	134514	CTGGCTCCAGGATATTTTTCAGGGCATGCTGAAGAG-----CAAAAGAGATGTACT	134567
QY	192	gggagatgaacaagagagaagaacaag-cccatgagctccctgtcaagaagagag	250
Db	134568	GGGAGATGAACACAGGAGGAAGAAAGCAAGCCCATGAAATTCCTGTCAAGAGGAAG	134627
QY	251	aacccccggaaaaaactgtgtgctggcagcagagaagaagatgltgaataattacatc	310
Db	134628	AACCTCCGTGAATAAACTGTTGACGTGGAGCGAGAAAGAAATGTGTAATAATTACTCTG	134687
QY	311	aaataccacagactagagaatgctgcagaagaaggctggaagatccaatgtaactgtgagct	370
Db	134688	AAATTCACACAGGAGAGAAAGAAATGCAGAGAGGGCTGAAACAAATCCATGTACCCGTGAGCT	134747
QY	371	tggagagtgaagaagactgctcggcagcctgagtttgggaattcttcgaattccaacaag	430
Db	134748	TGTAAAGTAAGAAACTGCTCCCGGACGAGCTAGGTTGGGACTTCTTCAGTTTCAACAAAG	134807
QY	431	gtctgcatctcgataaacaactatggttaacttctgataagcttgaagagaagactcaaa	490
Db	134808	GTCGTGCTATCTTAACACAAACCTATCATCAAGACCTTGATTAAGTTAAGGAAGAGCTCAAA	134867

QY	491	gatttggtttgaatcttccttcacatctccagaagaatctgaagaatgatggaactgga	550
Db	134868	GATTTGGTTTGAATGGCTCTTCAATCTCCAGAAAGTCTGAAGATGATGAGAAACTGMAA	134927
QY	551	agaaggaaggaagcatttgggattgtcacaaagtltcagcttggaaactgga---accagaag	607
Db	134928	GGAGGAAGGACTGATTTGGGATTAATACAGTGTTCAGCTGGAACCGAAGGAACACAGAGG	134987
QY	608	atacagagggcaagaagaaggaagaagcagagcgctttgggatttccctgataatgaatgc	667
Db	134988	ATACGAGAGCAAAAGAAATAGGAAAGAGCGCTCGCTATGGGATTCCTGCTGAAGATTC	135047
QY	668	ctgataacttctgtctctccagatgtttccatcttcctctctctcttcttcttggtaacata	727
Db	135048	ATGCTGCTTTCGTCTCCAGTGGTTTCCATTTCGTGATTTCTTCGCTCAAAATATA	135107
QY	728	ccataatgcacagatcatgtgctcctcagctcgtcgtcgcacatgagggagcatgaccacag	787
Db	135108	CCTAATGACAGACGTCACGACTTACTCTCTGCTTGCAATGAGGGACATGTTCCCAAC	135167
QY	788	tacatccatgaacatcgtgcgacagcttctactattgtcgtttcagcttaagttggttgt	847
Db	135168	TACATGTGGGAACCTCCACAGCAAAATTGACTATTGATGTTCTCAAC-TTAAAGCTGTGT	135226
QY	848	gtttttgttttgattatgttctgtctgtaataaaaaaataag	892
Db	135227	GTTTTGTTTTGATT-TTTTGTCTGTTRAGTTAAAAAATAAATAA	135270
RESULT	10		
LOCUS	AX210422	408 bp	DNA
DEFINITION	Sequence 64 from Patent WO0157058.		Linear
ACCESSION	AX210422		
VERSION	AX210422.1	GI:15424682	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 408)		
TITLE	Rosenthal, A., Hinzmann, B., Schaefer, R., Zuber, J., Tchernitsa, O., Grips, M., Hellriegel, M., Schmitz, A.C. and Sers, C.		
JOURNAL	Detection of differential gene expression		
FEATURES	Patent: WO 0157058-A 64 03-AUG-2001;		
source	Metagen Gesellschaft fuer Genomforschung mbH (DE)		
	Location/Qualifiers		
	1..408		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	141 a 111 g 88 t 2 others		
ORIGIN			
	37.4%; Score 334; DB 6; Length 408;		
Query Match	Best Local Similarity 98.0%; Pred. No. 1e-68; Mismatches 2; Indels 6; Gaps 6;		
Matches	400; Conservative 0;		
QY	190	ctgggagatgaaacagaggaagaagaacaaagccattgagcttccctgtcaagaaggaa	249
Db	1	CTGGAGATGGAACAGAGGAAGAAGAAACAAAGCCATTGAGCTCCCTGTCAAGAGAA	60
QY	250	gaacccccctgaaaaaacctgtgtgattgtgcgacgagagaagaaggtgtgaaaaattacatct	309
Db	61	GAAACCCCTGAAAAAAGCTGTGATGTGGACAGAGAAAGAAAGTGCTGAAAAATTACATCT	120
QY	310	gaatacacacagactgagagaaatgcagaagaagggctgtaacgaatcaatgacatgagc	369
Db	121	GAAATACCAACAGACTGGAAGAAATGCAGAAAGAGCGCTGAACGATTAANTGATACCTTTGAGC	180
QY	370	ttgagagtaagaaagctgtctcggcagctaggtttggatttcttcagtttccaaacaaa	429

Db	181	TTGGAGACTAGAAAGCTGCTCGGCGACAGTGGATTTCCTTCAGTTCACCAACAAA	240
Oy	430	ggtctgtcatctgatacaacaacctatggttcaactt-ggataagctgtaaggaaagact-c	487
Db	241	GCTGCTCATCTGTAAGAACAACTATGCTTAACTTGGGATTAAGCTGAAGAAAGAGCTTC	300
Oy	488	aaagtttggtttgatctctctctcaatctccaagaagtc-tgaagatgatga-gaaact	545
Db	301	AAAGTTTGGTTTGGTAATGCTCTTTCAATTCACGAAGAGCTTGAAGATGATGAGAAACT	360
Oy	546	gaaagaaga-ggaagagagcattt-ggagattgcacaagcttcagctcga	591
Db	361	GAAGAAAGAGGAGAGAGGAGGATTTGGGCAATTGTCACAACTTCAGCTGCA	408
RESULT 11			
AC073487/c			
LOCUS	AC073487	54666 bp	DNA linear PRI 03-NOV-2001
DEFINITION	Homo sapiens 12 BAC RP11-76217 (Roswell Park Cancer Institute Human		
ACCESSION	AC073487		
VERSION	AC073487.34	GI:14578058	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Ennaffia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
	1 (bases 1 to 54666)		
	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,		
	Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,I.,		
	Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Bouck,J.,		
	Boyle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Butlay,C.,		
	Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,		
	Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,		
	Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,		
	Cox,C., Coyle,M.D., Dabhorne,S.R., David,R., Davila,M.L., Davis,C.,		
	Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,		
	Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,		
	Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,		
	Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferriguto,D.,		
	Flahg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,		
	Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,		
	Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,		
	Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,		
	Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homs1,F.,		
	Howard,S., Huber,J., Huliy,K., Hume,J., Ioshikhes,I., Jackson,L.E.,		
	Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,		
	Karlsson,E., Kelly,S., Khan,U., King,L., Kovav,J., Kovar,C.,		
	Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,		
	Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,		
	Loulsaged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,		
	Ma,J., Maheshwari,M., Mapua,P., Maronde1,I., Martin,R.,		
	Mattindale,A., Martinez,E., Messay,E., Mawhney,E., McLeod,M.P.,		
	Meador,M., Mei,G., Mercher,S., Metzker,M., Miller,A., Miner,G.,		
	Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,		
	Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,		
	Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenko,S.,		
	Ogduh,N., Okwomu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,		
	Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,		
	Quiles,M., Ren,Y., Rives,M., Rojs,A., Rojundokan,I., Rolfe,M.,		
	Ruiz,S., Savery,G., Scheer,S., Scott,G., Shen,H., Shm,C.,		
	Shooshtari,N., Sisson,I., Sodergren,E., Somalet,T., Sparks,A.,		
	Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabori,P., Tamerisa,A.,		
	Tamerisa,K., Tang,H., Tansy,J., Taylor,C., Taylor,T., Telford,B.,		
	Thomas,N., Thoms,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D.,		
	Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,		
	Washington,C., Wallington,S., Williams,G., Williamson,A.,		
	Wlaczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,		
	Zorilla,S., Zuchertapfli,R., Weinstein,G. and Gibbs,R.		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 54666)		
REFERENCE			
AUTHORS	Worley,K.C.		

TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 54666)
AUTHORS	Morley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (30-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 54666)
AUTHORS	Morley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (03-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Jun 30, 2001 this sequence version replaced gi:14575758. INFORMATION: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a> or email <a href="mailto:gc-help@bcm.tmc.edu">gc-help@bcm.tmc.edu</a>

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

Location/Qualifiers

1..54666

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="12"

/clone="RP11-76217"

8..292

/rpl\_family="AluSp"

293..320

/rpl\_family="(CAAA)n"

337..454

/rpl\_family="Aluub"

456..734

/rpl\_family="AluSx"

complement(1548..1602)

/rpl\_family="I2"

complement(4116..4163)

/rpl\_family="MIR"





## TITLE

JOURNAL

## COMMENT

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rella, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (10-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center -----  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information -----  
Center project name: L18853  
Center clone name: 320\_D\_22  
-----

\* NOTE: This record contains 74 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 725: contig of 725 bp in length  
\* 726 825: gap of 100 bp  
\* 826 1534: contig of 709 bp in length  
\* 1535 1634: gap of 100 bp  
\* 1635 2359: contig of 725 bp in length  
\* 2360 2459: gap of 100 bp  
\* 2460 3181: contig of 722 bp in length  
\* 3182 3281: gap of 100 bp  
\* 3282 4002: contig of 721 bp in length  
\* 4003 4102: gap of 100 bp  
\* 4103 4803: contig of 701 bp in length  
\* 4804 4903: gap of 100 bp  
\* 4904 5614: contig of 711 bp in length  
\* 5615 5714: gap of 100 bp  
\* 5715 6430: contig of 716 bp in length  
\* 6431 6530: gap of 100 bp  
\* 6531 7244: contig of 714 bp in length  
\* 7245 7344: gap of 100 bp  
\* 7345 8073: contig of 729 bp in length  
\* 8074 8173: gap of 100 bp  
\* 8174 8861: contig of 688 bp in length  
\* 8862 8961: gap of 100 bp  
\* 8962 9688: contig of 727 bp in length  
\* 9689 9788: gap of 100 bp  
\* 9789 10488: contig of 700 bp in length  
\* 10489 10588: gap of 100 bp  
\* 10589 11303: contig of 717 bp in length  
\* 11306 11405: gap of 100 bp  
\* 11406 12123: contig of 718 bp in length  
\* 12124 12223: gap of 100 bp  
\* 12224 12945: contig of 722 bp in length  
\* 12946 13045: gap of 100 bp  
\* 13046 13750: contig of 705 bp in length  
\* 13751 13850: gap of 100 bp  
\* 13851 14546: contig of 696 bp in length  
\* 14547 14646: gap of 100 bp  
\* 14647 15369: contig of 723 bp in length  
\* 15370 15469: gap of 100 bp  
\* 15470 16131: contig of 662 bp in length

16132 16231: gap of 100 bp  
\* 16232 16951: contig of 720 bp in length  
\* 16952 17051: gap of 100 bp  
\* 17052 17775: contig of 724 bp in length  
\* 17776 17875: gap of 100 bp  
\* 17876 18605: contig of 730 bp in length  
\* 18606 18705: gap of 100 bp  
\* 18706 19411: contig of 706 bp in length  
\* 19412 19511: gap of 100 bp  
\* 19512 20235: contig of 724 bp in length  
\* 20236 20335: gap of 100 bp  
\* 20336 21061: contig of 726 bp in length  
\* 21062 21161: gap of 100 bp  
\* 21162 21867: contig of 706 bp in length  
\* 21868 21967: gap of 100 bp  
\* 21968 22678: contig of 711 bp in length  
\* 22679 22778: gap of 100 bp  
\* 22779 23504: contig of 726 bp in length  
\* 23505 23604: gap of 100 bp  
\* 23605 24328: contig of 724 bp in length  
\* 24329 24428: gap of 100 bp  
\* 24429 25156: contig of 728 bp in length  
\* 25157 25256: gap of 100 bp  
\* 25257 25982: contig of 726 bp in length  
\* 25983 26082: gap of 100 bp  
\* 26083 26790: contig of 708 bp in length  
\* 26791 26890: gap of 100 bp  
\* 26891 27616: contig of 726 bp in length  
\* 27617 27716: gap of 100 bp  
\* 27717 28517: contig of 801 bp in length  
\* 28518 28617: gap of 100 bp  
\* 28618 29357: contig of 740 bp in length  
\* 29358 29457: gap of 100 bp  
\* 29458 30185: contig of 728 bp in length  
\* 30186 30285: gap of 100 bp  
\* 30286 31014: contig of 729 bp in length  
\* 31015 31114: gap of 100 bp  
\* 31115 31818: contig of 704 bp in length  
\* 31819 31918: gap of 100 bp  
\* 31919 32637: contig of 719 bp in length  
\* 32638 32737: gap of 100 bp  
\* 32738 33471: contig of 734 bp in length  
\* 33472 33571: gap of 100 bp  
\* 33572 34306: contig of 735 bp in length  
\* 34307 34406: gap of 100 bp  
\* 34407 35116: contig of 710 bp in length  
\* 35117 35216: gap of 100 bp  
\* 35217 35946: contig of 730 bp in length  
\* 35947 36046: gap of 100 bp  
\* 36047 36789: contig of 743 bp in length  
\* 36790 36889: gap of 100 bp  
\* 36890 37605: contig of 716 bp in length  
\* 37606 37705: gap of 100 bp  
\* 37706 38428: contig of 723 bp in length  
\* 38429 38528: gap of 100 bp  
\* 38529 39248: contig of 720 bp in length  
\* 39249 39348: gap of 100 bp  
\* 39349 40083: contig of 735 bp in length  
\* 40084 40183: gap of 100 bp  
\* 40184 40888: contig of 705 bp in length  
\* 40889 40988: gap of 100 bp  
\* 40989 41715: contig of 727 bp in length  
\* 41716 41815: gap of 100 bp  
\* 41816 42549: contig of 734 bp in length  
\* 42550 42649: gap of 100 bp  
\* 42650 43356: contig of 707 bp in length  
\* 43357 43456: gap of 100 bp  
\* 43457 44158: contig of 702 bp in length  
\* 44159 44258: gap of 100 bp  
\* 44259 44977: contig of 719 bp in length  
\* 44978 45077: gap of 100 bp  
\* 45078 45805: contig of 728 bp in length  
\* 45806 45905: gap of 100 bp



DEFINITION	Blastocyst cDNA.
ACCESSION	E25874

548 GGTATTTTGGTTTAAATTATTTTGGCTTATTAAATAAAAA 586

RESULT 15  
AX210889 255 bp DNA linear PAT 31-AUG-2001  
LOCUS Sequence 531 from Patent WO0157058.  
DEFINITION AX210889  
ACCESSION AX210889  
VERSION AX210889.1 GI:15425149  
KEYWORDS  
SOURCE Rattus sp.  
ORGANISM Rattus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 255)  
REFERENCE  
AUTHORS Rosenthal, A., Hinemann, B., Schaefer, R., Zuber, J., Tchernitsa, O.,  
Grips, M., Hellriegel, M., Schmitz, A.C. and Sers, C.  
TITLE Detection of differential gene expression  
JOURNAL Patent: WO 0157058-A 531 09-AUG-2001.  
Metagen Gesellschaft fuer Genomforschung mbH (DE)  
FEATURES  
source location/Qualifiers  
1..255  
/organism="Rattus sp."  
/db\_xref="taxon:10118"  
BASE COUNT 89 a 40 c 69 g 57 t  
ORIGIN

Query Match 23.7%; Score 211.8; DB 6; Length 255;  
Best Local Similarity 89.4%; Pred. No. 8.3e-40;  
Matches 228; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 189 actggagatgaaacagaggaagaacaaagccattgagctccctgtcaagaaga 248  
|||||  
Db 1 ACTGGAGATGAACTGAGAGAAACCAAGCCTATAGAACTGCTGTTAAAGAGA 60  
OY 249 agaaccctgaaaaaactgtgtgtgtgagcagcagaagaagaagtgtgaaattacatc 308  
|||||  
Db 61 AGAACCTCCTGTAAGAGTTGTGATATGCAATCAGAAAAAGAGTGATAAAAATTACATC 120  
OY 309 tgaataccacagactgagagatgcagaaagagggctgaacgatccaatgtacctgtgag 368  
|||||  
Db 121 TGGATACCTCAACTAGAGAGATGCAGAGAGGGCTGAACCTTCAATGTGCTGTAG 180  
OY 369 ctggagagtaagaagctgtcggcagctaggtttggatttctcaagttccaacaaa 428  
|||||  
Db 181 CTTGGAGAGTAAGAGGCTGCTGGGCGAGCGAGGTTTGAATTCTCAGTTCACAACAA 240  
OY 429 aggtctgtcatctga 443  
|||||  
Db 241 AGGTTATCATCTCTGA 255

Search completed: June 19, 2002, 14:23:34  
Job time: 9027 sec